

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 29, 2003, 15:16:07 ; Search time 40 Seconds
(Without alignments)
194.672 Million cell updates/sec

Title: US-09-924-102-2
Perfect score: 418
Sequence: 1 MLSTHFLYFLFYFLSYSL.....RMGGGGRGGTADTGMEFLS 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_73:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	72	17.2	845	2 D96799	hypothetical prote
2	71	17.0	627	4 A40201	artifact-warning s
3	68	16.3	462	2 H64048	signal recognition
4	68	16.3	511	2 A56750	archaen - human
5	67	16.0	118	2 T15239	hypothetical prote
6	67	16.0	247	2 S78470	NBDH2 dehydrogenas
7	67	16.0	829	2 S72366	DNA topoisomerase
8	67	16.0	1145	2 T18235	transcription acti
9	65.5	15.7	613	4 C40201	artifact-warning s
10	65	15.6	80	2 T10550	hypothetical prote
11	65	15.6	326	2 A41732	heterogeneous ribo
12	65	15.6	333	2 T52355	hypothetical prote
13	65	15.6	386	1 S22315	snRNP-associated p
14	65	15.6	719	2 T52510	hypothetical prote
15	64.5	15.4	129	2 S14964	glycine-rich prote
16	64.5	15.4	343	2 S55659	legumetin protein 6
17	64	15.3	131	2 G90335	conserved hypotet
18	64	15.3	330	2 T25940	hypothetical prote
19	64	15.3	765	1 TSHU1	DNA topoisomerase
20	64	15.3	767	2 J00144	DNA topoisomerase
21	64	15.3	767	2 A49546	DNA topoisomerase
22	64	15.3	1306	2 A70934	hypothetical glyci
23	63.5	15.2	389	2 T44957	heat shock protein
24	63.5	15.2	391	2 E84207	hypothetical glyci
25	63.5	15.2	749	2 A70812	hypothetical prote
26	63	15.1	163	2 T23076	hypothetical glyci
27	63	15.1	432	2 B84620	hypothetical prote
28	63	15.1	439	2 D70954	hypothetical glyci
29	63	15.1	476	2 AG1031	hypothetical prote

30	63	15.1	1079	2 B70807	hypothetical glyci
31	62.5	15.0	276	2 T33925	hypothetical prote
32	62.5	15.0	463	2 T36810	probable integral
33	62.5	15.0	1410	1 A57013	early endosome ant
34	62	14.8	161	2 G71407	transcription fact
35	62	14.8	167	2 T20548	hypothetical prote
36	62	14.8	180	2 T28938	hypothetical prote
37	62	14.8	314	2 T08675	hypothetical prote
38	62	14.8	1898	1 A45973	trichohyalin - hum
39	61.5	14.7	268	2 S09860	hypothetical prote
40	61.5	14.7	272	2 T34184	hypothetical prote
41	61.5	14.7	284	2 T23158	hypothetical prote
42	61.5	14.7	806	2 S22765	heterogeneous ribo
43	61	14.6	78	2 B54897	dermatopilin b I pr
44	61	14.6	155	2 C86206	hypothetical prote
45	61	14.6	271	2 S34666	glycine-rich prote

ALIGNMENTS

RESULT 1
D96799
hypothetical protein F22K20.13 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96799
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maltli, R.; Marzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96799
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-845 <STO>
A:Cross-references: GB:AE051173; NID:92829912; PIDN:AMC00620.1; GSPDB:GN00141
C:Genetics:
A:Gene: F22K20.13
A:Map position: 1
Query Match
Best Local Similarity 39.2%; Score 72; DB 2; Length 845;
Matches 20; Conservative 9; Mismatches 18; Indels 4; Gaps 2;
OY 28 LRKTKQOKKEQIIROSEVLEFRSETLKRKTKRKGCG-OGGRCGTADTGG 77
Db 748 VRKEROOKANKVSYLSK---RGGRGGRGARGGARGGARGGSGRDFG 795
RESULT 2
A40201
artifact-warning sequence (translated ALU class A) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C:Accession: A40201
R:Claverie, J.M.
Personal communication, 1992
A:Reference number: A40201
A:Accession: A40201
A:Molecule type: DNA
A:Residues: 1-627 <CLA>
R:Claverie, J.M.
Genomics 12, 838-841, 1992
A:Title: Identifying coding exons by similarity search: Alu-derived and other potenti
A:Reference number: A40200; MUID:9241891; PMID:1572661
A:Contents: annotation

Db 216 TISPGHSHRR-SFSSGAGGPPD 237

RESULT 7

S72366

DNA topoisomerase (EC 5.99.1.2) I, somatic - African clawed frog

A:Species: Xenopus laevis (African clawed frog)

C:Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 16-Jul-1999

C:Accession: S72366

R:Pandit, S.D.; Richard, R.E.; Sternglanz, R.; Bogenhagen, D.F.

Nucleic Acids Res. 24, 3593-3600, 1996

A:Title: Cloning and characterization of the gene for the somatic form of DNA topoisomerase

A:Reference number: S72366; MUID:96433160; PMID:8836188

A:Accession: S72366

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-829 <P>N>

A:Cross-references: EMBL:L07777; NID:9214833; PIDN:AMB36608.1; PID:9214834

A:Experimental source: somatic cells

C:Comment: The type I DNA topoisomerase catalyzes the ATP-independent transient breakage

reak in another, followed by rejoining. This reaction will lead to the conversion of one

C:Genetics:

A:Gene: Top1

C:Superfamily: eukaryotic type I DNA topoisomerase

C:Keywords: DNA binding; DNA replication; isomerase

F:779/Active site: Tyr #status predicted

Query Match

Best Local Similarity 16.0%; Score 67; DB 2; Length 829;

Matches 18; Conservative 9; Mismatches 33; Indels 0; Gaps 0;

QY 18 YSLGDRARLCLRTKQOQKEQILRQSEVLFSETLRTKTKGKGRMGCGGGGTADTGC 77

Db 500 YETARRLKMCEKERNYKEMKSKEMKVRRAVALTIDKLALRAGNEKEGETADTVC 559

RESULT 8

T18235

transcription activator GAL1 homolog - yeast (Candida albicans)

C:Species: Candida albicans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T18235

R:Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, November 1998

A:Reference number: Z18831

A:Accession: T18235

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1145 <BAR>

A:Cross-references: EMBL:AL033501; NID:e1341022; PID:e1341031; PIDN:CAA21993.1

C:Genetics:

A:Note: CA41C10.09

Query Match 16.0%; Score 67; DB 2; Length 1145;

Best Local Similarity 30.4%; Pred. No. 32;

Matches 21; Conservative 15; Mismatches 23; Indels 10; Gaps 3;

QY 16 LYSISGDRARLCLRTKQOQKEQILRQSEVLFSETLRTKTKGKGRMGCGGGR----- 69

Db 842 IQHMSGPGQVYLVOQOQOQOQOQPOQOQSVSASQSKSVAKTKKGT---GQGRKKRASI 898

QY 70 -GGTADTGC 77

Db 899 SAGTAPTPG 907

RESULT 9

C40201

artifact-warning sequence (translated ALT class C) - human

C:Species: Homo sapiens (man)

C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000

C:Accession: C40201

R:Claverie, J.M.

personal communication, 1992

A:Reference number: A40201

A:Accession: C40201

A:Molecule type: DNA

A:Residues: 1-613 <CLA>

R:Claverie, J.M.

Genomics 12, 838-841, 1992

A:Title: Identifying coding exons by similarity search: Alu-derived and other potentl

A:Reference number: A40200; MUID:92241891; PMID:1572661

A:Contents: annotation

C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames o

in-frame stop codons are shown as 'x'.

C:Comment: Any significant similarity of a predicted protein sequence to a portion of

Query Match 15.7%; Score 65.5; DB 4; Length 613;

Best Local Similarity 30.1%; Pred. No. 26;

Matches 22; Conservative 5; Mismatches 15; Indels 31; Gaps 3;

QY 16 LYSISGDRARLCLRTKQOQKEQILRQSEVLFSETLRTKTKGK-----RRMGSG 67

Db 184 LHSSIGDVRVRLCLTKKKK-----KXXXXGAAVAHACNPSTLGGKG 223

QY 68 G---RGGTADTGC 77

Db 224 GWMRPGVRDQPG 236

RESULT 10

T10550

hypothetical protein T12613.70 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999

C:Accession: T10550

R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Ban

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16533

A:Accession: T10550

A:Molecule type: DNA

A:Residues: 1-80 <BEV>

A:Cross-references: EMBL:AL080252; GSPDB:GN00062; ATSP:T12613.70

A:Experimental source: cultivar Columbia; BAC clone T12613

C:Genetics:

A:Gene: ATSP:T12613.70

A:Map position: 4

Query Match 15.6%; Score 65; DB 2; Length 80;

Best Local Similarity 41.9%; Pred. No. 4.2;

Matches 13; Conservative 5; Mismatches 11; Indels 2; Gaps 1;

QY 49 RSETLRKTKGKGRMG--QGGRGCTADTGC 77

Db 3 KSDAYRRKSGKPNKRWGGMGGGGGGGSGG 33

RESULT 11

A41732

heterogeneous ribonuclear particle protein hrp36 - fruit fly (Drosophila melanogaster

N:Alternate names: heterogeneous nuclear RNP protein hrp36

C:Species: Drosophila melanogaster

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-Jul-1999

C:Accession: A41732

R:Matunis, E.L.; Matunis, M.J.; Dreyfuss, G.

J. Cell Biol. 116, 257-269, 1992

A:Title: Characterization of the major hnRNP proteins from Drosophila melanogaster.

A:Reference number: A41732; MUID:92112968; PMID:1730754

A:Accession: A41732

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-326 <MAT>

A:Cross-references: GB:X62636; NID:g11035; PIDN:CAA44502.1; PID:g11036

A:Note: sequence extracted from NCBI backbone (NCBIN:76634, NCBIIP:76635)

C:Genetics:

A:Gene: FlyBase:Hrb87F
 A:Cross-references: FlyBase:FBgn0004237
 C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology
 C:Keywords: DNA binding
 F:25-91/Domain: ribonucleoprotein repeat homology <RKM1>
 F:116-182/Domain: ribonucleoprotein repeat homology <RKM2>

Query Match 15.6%; Score 65; DB 2; Length 326;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 13; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 54 RKTGKGRMGCGGGRGTADTGMP 79
 Db 231 RQNG--GGMWGGAGGCGFGNGGNF 254

RESULT 12

T52355
 Hypothetical protein B11E6.40 [Imported] - Neurospora crassa

C:Species: Neurospora crassa
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: T52355
 R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, September 2000
 A:Reference number: Z26053
 A:Accession: T52355
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-333 <SCH>

A:Cross-references: EMBL:AL442043; GSPDB:GN00116; NCSP:B11E6.40
 A:Experimental source: BAC clone B11E6; strain OR/4A
 C:Genetics:
 A:Gene: NCSP:B11E6.40
 A:Map position: 6

Query Match 15.6%; Score 65; DB 2; Length 333;
 Best Local Similarity 52.6%; Pred. No. 16;
 Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 59 KGRMGCGGGRGTADTG 77
 Db 232 QGEQMGESGGRGTGAGG 250

RESULT 13

S22315
 snRNP-associated protein P11 - fruit fly (Drosophila melanogaster)
 N:Alternate names: heterogeneous ribonuclear particle (hnRNP) protein A1 homolog; hnRNP C;Species: Drosophila melanogaster
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S22315; S65539; S13731

R:Hovemann, B.T.; Dessen, E.; Mechler, H.; Mack, E.
 Nucleic Acids Res. 19, 4909-4914, 1991
 A:Title: Drosophila snRNP associated protein P11 which specifically binds to heat shock
 A:Reference number: S22315; M01D:92020124; PMID:1717937

A:Accession: S22315
 A:Molecule type: DNA
 A:Residues: 1-386 <HOV>
 A:Cross-references: EMBL:X59691

R:Reiner, B.
 submitted to the EMBL Data Library, August 1991
 A:Reference number: S65539
 A:Accession: S65539

A:Molecule type: DNA
 A:Residues: 1-270, 'T', 272-386 <REI>
 A:Cross-references: EMBL:X59691; NID:g8317; PIDN:CAA42212.1; PID:g8318
 N:Haynes, S.R.; Johnson, D.; Raychaudhuri, G.; Beyer, A.L.
 Nucleic Acids Res. 19, 25-31, 1991

A:Title: The Drosophila Hrb87F gene encodes a new member of the A and B hnRNP protein group
 A:Reference number: S13731; M01D:91187645; PMID:1849257
 A:Accession: S13731
 A:Status: preliminary
 A:Molecule type: mRNA

A:Residues: 1-386 <HAY>
 A:Cross-references: EMBL:X54803; NID:g8093; PIDN:CAA38574.1; PID:g8094
 C:Genetics:
 A:Gene: FlyBase:Hrb87F
 A:Cross-references: FlyBase:FBgn0004237
 A:Introns: 16/3

C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology
 F:25-91/Domain: ribonucleoprotein repeat homology <RKM1>
 F:116-182/Domain: ribonucleoprotein repeat homology <RKM2>

Query Match 15.6%; Score 65; DB 1; Length 386;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 13; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 54 RKTGKGRMGCGGGRGTADTGMP 79
 Db 231 RQNG--GGMWGGAGGCGFGNGGNF 254

RESULT 14

T52510
 Hypothetical protein B2J23.60 [Imported] - Neurospora crassa

C:Species: Neurospora crassa
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: T52510
 R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, September 2000
 A:Reference number: Z26053
 A:Accession: T52510
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-719 <SCH>

A:Cross-references: EMBL:AL442164; GSPDB:GN00116; NCSP:B2J23.60
 A:Experimental source: BAC clone B2J23; strain OR/4A
 C:Genetics:
 A:Gene: NCSP:B2J23.60
 A:Map position: 6
 A:Introns: 349/1; 601/1

Query Match 15.6%; Score 65; DB 2; Length 719;
 Best Local Similarity 37.8%; Pred. No. 34;
 Matches 17; Conservative 5; Mismatches 17; Indels 6; Gaps 2;

QY 38 QQLRQSEVLFNSETL--RKTGKGRMGCGGGR--GGTADTG 76
 Db 298 EDVLEQGQDLFASQDVIADRTVGHSDGRGAGAGNIVAGAGDGG 342

RESULT 15

S14984
 glycine-rich protein (clones uA-3 and uK-4) - tomato (fragment)
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 13-Jan-1995 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
 C:Accession: S14984; S14985

R:Howalter, A.M.; Zhou, J.; Rumeau, D.; Worst, S.G.; Varner, J.E.
 Plant Mol. Biol. 16, 547-565, 1991
 A:Title: Tomato extensin and extensin-like CDNs: structure and expression in response to wounding
 A:Reference number: S14970; M01D:91329690; PMID:1714316

A:Accession: S14984

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-129 <SHO>
 A:Cross-references: EMBL:X55695; NID:g19222; PIDN:CAA39224.1; PID:g388258
 A:Experimental source: cv. UC82B; haplotype n12; stem tissue; clone uA-3

A:Accession: S14985
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 71-129 <SHO>

A:Cross-references: EMBL:X55696; NID:g19223; PIDN:CAA39224.1; PID:g1345534
 A:Experimental source: cv. UC82B; haplotype n12; stem tissue; clone uK-4
 Query Match 15.4%; Score 64.5; DB 2; Length 129;
 Best Local Similarity 58.3%; Pred. No. 7.5;

	Matches	14;	Conservative	3;	Mismatches	6;	Indels	1;	Gaps	1;
Oy	57	GKGRBWMGO-GGGRGTADTGGMF	79							
Db	52	GGGGRGTGGRSGGGRGGRDSSGRF	75							

Search completed: May 29, 2003, 15:21:03
 Job time : 42 secs

THIS PAGE BLANK (USPTO)